



Translation of the front page of the priority document of
Japanese Patent Application No.11-220496)

PATENT OFFICE
JAPANESE GOVERNMENT

RECEIVED
APR 09 2001
Technology Center 2001

This is to certify that the annexed is a true copy of the
following application as filed with this Office.

Date of Application: August 3, 1999

Application Number : Patent Application 11-220496

Applicant(s) : Canon Kabushiki Kaisha

August 25, 2000

Commissioner,
Patent Office

Kouzo OIKAWA

Certification Number 2000-3067804



日本国特許庁
PATENT OFFICE
JAPANESE GOVERNMENT

CFM 1969 US
P200-0123

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日

Date of Application:

1999年 8月 3日

出願番号

Application Number:

平成11年特許願第220496号

出願人

Applicant(s):

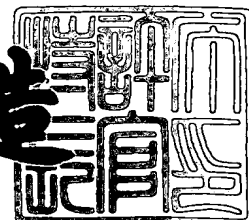
キヤノン株式会社

CERTIFIED COPY OF
PRIORITY DOCUMENT

2000年 8月25日

特許庁長官
Commissioner,
Patent Office

及川耕造



出証番号 出証特2000-3067804

【書類名】 特許願

【整理番号】 3937009

【提出日】 平成11年 8月 3日

【あて先】 特許庁長官殿

【国際特許分類】 G10L 3/00

【発明の名称】 音声合成用の素片辞書作成方法及び前記素片辞書を用いた音声合成方法及び装置

【請求項の数】 24

【発明者】

【住所又は居所】 東京都大田区下丸子3丁目30番2号 キヤノン株式会社
社内

【氏名】 山田 雅章

【特許出願人】

【識別番号】 000001007

【氏名又は名称】 キヤノン株式会社

【代理人】

【識別番号】 100076428

【弁理士】

【氏名又は名称】 大塚 康德

【電話番号】 03-5276-3241

【選任した代理人】

【識別番号】 100093908

【弁理士】

【氏名又は名称】 松本 研一

【電話番号】 03-5276-3241

【選任した代理人】

【識別番号】 100101306

【弁理士】

【氏名又は名称】 丸山 幸雄

【電話番号】 03-5276-3241

【手数料の表示】

【予納台帳番号】 003458

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9704672

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 音声合成用の素片辞書作成方法及び前記素片辞書を用いた音声合成方法及び装置

【特許請求の範囲】

【請求項 1】 音声が記録されたデータベースから合成単位に切り出された素片データを読み込む読み込みステップと、

前記読み込みステップで読み込まれた素片データに対して符号化方式を決定する符号化方式決定ステップと、

前記符号化方式決定ステップで決定された符号化方式に対する復号化方法を指示する情報を素片辞書に記録する復号化情報記録ステップと、

前記符号化方式決定ステップで決定された前記符号化方式に従って前記読み込みステップで読み込まれた前記素片データを符号化する符号化ステップと、

前記符号化ステップで符号化された素片データを前記素片辞書に記録する符号化データ記録ステップと、

を有することを特徴とする音声合成用の素片辞書作成方法。

【請求項 2】 前記符号化方式決定ステップで決定される符号化方式に応じて、前記素片データに割り当てるビット数を変更することを特徴とする請求項 1 に記載の素片辞書作成方法。

【請求項 3】 更に、複数の量子化テーブルを作成する量子化テーブル作成ステップを設け、

前記符号化方式決定ステップにおいて前記量子化テーブル作成ステップで作成された量子化テーブルの内のいずれを用いるかを決定することを特徴とする請求項 1 又は 2 に記載の素片辞書作成方法。

【請求項 4】 音声の記録されたデータベースから合成単位に切り出された素片データを読み込む読み込みステップと、

前記読み込みステップで読み込まれた前記素片データに対して線形予測係数及び予測残差を求める線形予測ステップと、

前記線形予測ステップで求められた予測係数を記録する予測係数記録ステップと、

前記線形予測ステップで求められた予測残差を量子化する量子化方式を決定する残差量子化方式決定ステップと、

前記量子化方式を指示する情報を記録する残差量子化方式記録ステップと、

前記線形予測ステップで求められた線形予測係数及び前記残差量子化方式決定ステップで決定された量子化方式に基づいて前記読み込みステップで読み込まれた前記素片データを予測符号化する予測符号化ステップと、

前記予測符号化ステップで符号化された素片データを前記素片辞書に記録する符号化データ記録ステップと、

を有することを特徴とする音声合成用の素片辞書作成方法。

【請求項 5】 音声合成用の素片データが記録された音声合成用素片辞書における、所望の素片データが記録された位置を得る素片位置検索ステップと、

素片データを復号化するための情報を読み込む復号化情報読み込みステップと、

前記素片データが記憶された位置に基づいて、符号化された素片データを読み込む符号化データ読み込みステップと、

前記読み込みステップで読み込まれた符号化データを前記復号化情報読み込みステップで読み込まれた情報に従って復号する復号ステップと、

前記復号ステップで復号された素片データを接続して音声を合成する音声出力ステップと、

を有することを特徴とする音声合成方法。

【請求項 6】 前記復号化するための情報は、前記素片データに割り当てられているビット数であることを特徴とする請求項 5 に記載の音声合成方法。

【請求項 7】 前記復号化するための情報は、復号に用いる量子化テーブルに関する情報であることを特徴とする請求項 5 又は 6 に記載の音声合成方法。

【請求項 8】 音声合成用の素片データが記録された音声合成用素片辞書における、所望の素片データが記録された位置を得る素片位置検索ステップと、

前記素片データに対応する線形予測係数を読み込む予測係数読み込みステップと、

前記素片データに対応する予測残差復号方式を得る予測残差復号方式取得ステップと、

予測符号化された素片データを読み込む予測符号化データ読み込みステップと、

前記読込みステップで読込まれた予測符号化データを前記線形予測係数および前記予測残差復号方式に従って復号する予測復号ステップと、

前記予測復号ステップで予測復号された素片データを接続して音声を合成する音声出力ステップと、
を有することを特徴とする音声合成方法。

【請求項 9】 音声合成用の素片データが記録された音声合成用素片辞書における、所望の素片データが記録された位置を求める素片位置検索手段と、

素片データを復号化するための情報を求め、前記素片辞書から読み出された符号化された素片データを前記情報に従って復号する復号手段と、

前記復号手段で復号された素片データを接続して音声を合成する音声出力手段と、
を有することを特徴とする音声合成装置。

【請求項 1 0】 前記復号化するための情報は、前記素片データに割り当てられているビット数であることを特徴とする請求項 8 に記載の音声合成装置。

【請求項 1 1】 前記復号化するための情報は、復号に用いる量子化テーブルに関する情報であることを特徴とする請求項 9 又は 1 0 に記載の音声合成装置。

【請求項 1 2】 音声合成用の素片データが記録された音声合成用素片辞書における、所望の素片データが記録された位置を求める素片位置検索手段と、

前記素片データに対応する線形予測係数を求める予測係数判定手段と、

前記素片データに対応する予測残差復号方式を得る予測残差復号方式取得手段と、

前記素片辞書に格納された予測符号化された素片データを読込む予測符号化データ読込み手段と、

前記読込み手段で読込まれた予測符号化データを前記線形予測係数および前記予測残差復号方式に従って復号する予測復号手段と、

前記予測復号手段で予測復号された素片データを接続して音声を合成する音声出力手段と、
を有することを特徴とする音声合成装置。

【請求項 1 3】 音声合成用の素片辞書を作成する処理を実行するプログラムを記憶したコンピュータにより読取り可能な記憶媒体であって、

音声記録されたデータベースから合成単位に切り出された素片データを読み込む読み込みステップモジュールと、

読み込みステップで読み込まれた素片データに対して符号化方式を決定する符号化方式決定ステップモジュールと、

符号化方式決定ステップで決定された符号化方式に対する復号化方法を指示する情報を素片辞書に記録する復号化情報記録ステップモジュールと、

符号化方式決定ステップで決定された前記符号化方式に従って読み込みステップで読み込まれた前記素片データを符号化する符号化ステップと、

符号化ステップで符号化された素片データを前記素片辞書に記録する符号化データ記録ステップモジュールと、
を有する記憶媒体。

【請求項 1 4】 音声合成用の素片辞書を作成する処理を実行するプログラムを記憶したコンピュータにより読取り可能な記憶媒体であって、

音声の記録されたデータベースから合成単位に切り出された素片データを読み込む読み込みステップモジュールと、

読み込みステップで読み込まれた前記素片データに対して線形予測係数及び予測残差を求める線形予測ステップモジュールと、

線形予測ステップで求められた予測係数を記録する予測係数記録ステップモジュールと、

線形予測ステップで求められた予測残差を量子化する量子化方式を決定する残差量子化方式決定ステップモジュールと、

前記量子化方式を指示する情報を記録する残差量子化方式記録ステップモジュールと、

線形予測ステップで求められた線形予測係数及び残差量子化方式決定ステップで決定された量子化方式に基づいて読み込みステップで読み込まれた前記素片データを予測符号化する予測符号化ステップモジュールと、

予測符号化ステップで符号化された素片データを前記素片辞書に記録する符号

化データ記録ステップモジュールと、
を有することを特徴とする記憶媒体。

【請求項 1 5】 音声合成用の素片辞書を用いて音声を作成して出力する処理を実行するプログラムを記憶したコンピュータにより読取り可能な記憶媒体であって、

音声合成用の素片データが記録された音声合成用素片辞書における、所望の素片データが記録された位置を得る素片位置検索ステップモジュールと、

素片データを復号化するための情報を読込む復号化情報読込みステップモジュールと、

前記素片データが記憶された位置に基づいて、符号化された素片データを読込む符号化データ読込みステップモジュールと、

前記読込みステップで読込まれた符号化データを前記復号化情報読込みステップで読込まれた情報に従って復号する復号ステップモジュールと、

前記復号ステップで復号された素片データを接続して音声を作成する音声出力ステップモジュールと、

を有することを特徴とする記憶媒体。

【請求項 1 6】 音声合成用の素片辞書を用いて音声を作成して出力する処理を実行するプログラムを記憶したコンピュータにより読取り可能な記憶媒体であって、

音声合成用の素片データが記録された音声合成用素片辞書における、所望の素片データが記録された位置を得る素片位置検索ステップモジュールと、

前記素片データに対応する線形予測係数を読込む予測係数読込みステップモジュールと、

前記素片データに対応する予測残差復号方式を得る予測残差復号方式取得ステップモジュールと、

予測符号化された素片データを読込む予測符号化データ読込みステップモジュールと、

読込みステップで読込まれた予測符号化データを前記線形予測係数および前記予測残差復号方式に従って復号する予測復号ステップモジュールと、

予測復号ステップで予測復号された素片データを接続して音声を合成する音声出力ステップモジュールと、

を有することを特徴とする音声合成方法。

【請求項 1 7】 音声記録されたデータベースから合成単位に切り出された素片データを読込む読込み手段と、

前記読込み手段によって読み込まれた素片データに対して符号化方式を決定する符号化方式決定手段と、

前記符号化方式決定手段によって決定された符号化方式の対する復号化方法を指示する情報を素片辞書に記録する復号化情報記録手段と、

前記符号化方式決定手段によって決定された前記符号化方式に従って前記読込み手段によって読み込まれた前記素片データを符号化する符号化手段と、

前記符号化手段で符号化された素片データを前記素片辞書に記録する符号化データ記録手段と、

を備えたことを特徴とする音声合成用の素片辞書作成装置。

【請求項 1 8】 前記符号化方式決定手段によって決定される符号化方式に応じて、素片データに割り当てるビット数を変更するように構成したことを特徴とする請求項 1 7 に記載の素片辞書作成装置。

【請求項 1 9】 前記符号化決定手段は、複数の量子化テーブルを有し、前記量子化テーブルの内のいずれかを選択して決定することを特徴とする請求項 1 7 又は 1 8 に記載の素片辞書作成装置。

【請求項 2 0】 音声の記録されたデータベースから合成単位に切り出された素片データを読込む読込み手段と、

前記読込み手段によって読み込まれた前記素片データに対して線形予測係数及び予測残差を求める線形予測手段と、

前記線形予測手段によって求められた予測係数を記録する予測係数記録手段と、

前記線形予測手段によって求められた予測残差を量子化する量子化方式を決定する残差量子化方式決定手段と、

前記量子化方式を指示する情報を記録する残差量子化方式記録手段と、

前記線形予測手段で求められた線形予測係数及び前記残差量子化方式決定手段によって決定された量子化方式に基いて、前記読み込み手段によって読み込まれた前記素片データを予測符号化する予測符号化手段と、

前記予測符号化手段によって符号化された素片データを前記素片辞書に記録する符号化データ記録手段と、

を備えたことを特徴とする音声合成用の素片辞書作成装置。

【請求項 2 1】 音声の記録されたデータベースから合成単位に切り出された素片データを読み込む読み込み手段と、

前記読み込み手段によって読み込まれた前記素片データに対して声道フィルタ係数及び残差を求めるモデルパラメータ演算手段と、

前記モデルパラメータ演算手段によって求められた声道フィルタ係数を記録する係数記録手段と、

前記モデルパラメータ演算手段によって求められた残差を量子化する量子化方式を決定する残差量子化方式決定手段と、

前記量子化方式を指示する情報を記録する残差量子化方式記録手段と、

前記モデルパラメータ演算手段によって求められた声道フィルタ係数及び前記残差量子化方式決定手段によって決定された量子化方式に基いて、前記読み込み手段によって読み込まれた前記素片データを符号化する符号化手段と、

前記符号化手段によって符号化された素片データを前記素片辞書に記録する符号化データ記録手段と、

を備えたことを特徴とする音声合成用の素片辞書作成装置。

【請求項 2 2】 音声合成用の素片データが記録された音声合成用素片辞書における、所望の素片データが記録された位置を得る素片位置検索手段と、

前記素片データに対応する声道フィルタ係数を読み込む係数読み込み手段と、

前記素片データに対応する残差復号方式を得る残差復号方式取得手段と、

符号化された素片データを読み込む符号化データ読み込み手段と、

前記読み込み手段によって読み込まれた符号化データを前記声道フィルタ係数および前記残差復号方式に従って復号する復号手段と、

前記復号された素片データを接続して音声を合成する音声出力手段と、

を有することを特徴とする音声合成方法。

【請求項 2 3】 音声の記録されたデータベースから合成単位に切り出された素片データを読み込む読み込みステップと、

前記読み込みステップで読み込まれた前記素片データに対して声道フィルタ係数及び残差を求めるモデルパラメータ演算ステップと、

前記モデルパラメータ演算ステップで求められた声道フィルタ係数を記録する係数記録ステップと、

前記モデルパラメータ演算ステップで求められた残差を量子化する量子化方式を決定する残差量子化方式決定ステップと、

前記量子化方式を指示する情報を記録する残差量子化方式記録ステップと、

前記モデルパラメータ演算ステップで求められた声道フィルタ係数及び前記残差量子化方式決定手段によって決定された量子化方式に基づいて、前記読み込みステップで読み込まれた前記素片データを符号化する符号化ステップと、

前記符号化手段ステップで符号化された素片データを前記素片辞書に記録する符号化データ記録ステップと、

を有することを特徴とする音声合成用の素片辞書作成方法。

【請求項 2 4】 音声合成用の素片データが記録された音声合成用素片辞書における、所望の素片データが記録された位置を得る素片位置検索ステップと、

前記素片データに対応する声道フィルタ係数を読み込む係数読み込みステップと、

前記素片データに対応する残差復号方式を得る残差復号方式取得ステップと、

符号化された素片データを読み込む符号化データ読み込みステップと、

前記読み込みステップで読み込まれた符号化データを前記声道フィルタ係数および前記残差復号方式に従って復号する復号ステップと、

前記復号された素片データを接続して音声を合成する音声出力ステップと、

を有することを特徴とする音声合成方法。

【発明の詳細な説明】

【0 0 0 1】

【発明の属する技術分野】

本発明は、素片辞書を用いて音声を合成する音声合成方法及び装置に関するも

のである。

【0002】

【従来の技術】

計算機を用いて音声合成する音声合成技術では素片辞書を用いている。この素片辞書は、音声波形或は音声をパラメータに変換したデータを記録したものであり、音素、CV/VC、VCVといった適当な大きさ（合成単位）で前記データを検索できるようにしたものである。そして音声合成時には、その素片辞書から適切な素片を選択して抽出し、これら選択、抽出された素片を変形、接続することによって、所望の合成音声を得ている。これを説明するフローチャートが図2のフローチャートである。

【0003】

まずステップS1で、仮名漢字混じり文などで表現された発声内容を入力する。次にステップS2で、その入力された発声内容を解析し、素片記号列{p0, p1, …} および韻律決定のためのパラメータを得る。そして次にステップS3に進み、素片時間長、基本周波数、パワーといった韻律を決定する。次に素片辞書検索ステップS4において、ステップS2の入力解析で得られた素片記号列{p0, p1, …}、及びステップS3の韻律決定で得られた韻律に対して適切な素片{w0, w1, …}を素片辞書から検索する。そして次にステップS5に進み、ステップS4の素片辞書検索で得られた素片{w0, w1, …}を、ステップS3で決定された韻律に合うように変形して接続する。こうしてステップS6に進み、ステップS5における素片変形、接続処理の結果を基に、合成音声として出力する。

【0004】

また音声合成における有効な手法として、波形処理に基く合成方法がある。これは、声帯振動に同期して、波形重畳、ピッチ間隔の変更などの処理を施すものであり、合成音が自然に近く、かつ処理量の少なさを利点とする。このような波形処理に基く方法を用いた場合、素片辞書は、検索のためのインデックス、各素片に対応する波形データ、及びその他の情報から構成される。この際、波形データは、 μ -lawやADPCMといった手法を用いて圧縮されて素片辞書に格納され

る場合が多い。

【0 0 0 5】

【発明が解決しようとする課題】

上記従来例では、以下の問題点があった。

【0 0 0 6】

まず、圧縮方法として μ -law や A-law といった固定の量子化テーブルを用いて波形データを非一様に量子化した場合、十分な圧縮率を得ることが出来ないという問題がある。これは、あらゆる音声に対して最低限の品質を維持できるように量子化テーブルを設計する必要があるためである。

【0 0 0 7】

一方、ADPCM に代表される後向き適応予測、後向き適応量子化を用いた場合は、復号器の適応に要する演算量が問題となる。ここで、従来技術で述べたように、波形処理に基く音声合成は処理量の少なさが利点の一つであり、復号化に多くの演算を要するのでは意味がないことになる。

【0 0 0 8】

一般に、 μ -law や A-law といった固定的な符号化方式を用いると十分な圧縮性能を得ることができない。一方、入力に対して特性が変動する圧縮方式では適応的アルゴリズムが用いられることとなり、その演算量が問題となる。

【0 0 0 9】

本発明は上記従来例に鑑みてなされたもので、演算量を少なく抑え、かつ十分な圧縮性能が得られる素片データの符号化を行って作成された音声合成用の素片辞書作成方法及びその素片辞書を用いた音声合成方法及びその装置を提供することを目的とする。

【0 0 1 0】

【課題を解決するための手段】

上記目的を達成するために本発明の音声合成用の素片辞書の作成方法は以下のような工程を備える。即ち、

音声が記録されたデータベースから合成単位に切り出された素片データを読込む読込みステップと、

前記読込みステップで読込まれた素片データに対して符号化方式を決定する符号化方式決定ステップと、

前記符号化方式決定ステップで決定された符号化方式に対する復号化方法を指示する情報を素片辞書に記録する復号化情報記録ステップと、

前記符号化方式決定ステップで決定された前記符号化方式に従って前記読込みステップで読込まれた前記素片データを符号化する符号化ステップと、

前記符号化ステップで符号化された素片データを前記素片辞書に記録する符号化データ記録ステップと、

を有することを特徴とする。

また本発明の音声合成用の素片辞書作成方法は、

音声の記録されたデータベースから合成単位に切り出された素片データを読込む読込みステップと、

前記読込みステップで読み込まれた前記素片データに対して声道フィルタ係数及び残差を求めるモデルパラメータ演算ステップと、

前記モデルパラメータ演算ステップで求められた声道フィルタ係数を記録する係数記録ステップと、

前記モデルパラメータ演算ステップで求められた残差を量子化する量子化方式を決定する残差量子化方式決定ステップと、

前記量子化方式を指示する情報を記録する残差量子化方式記録ステップと、

前記モデルパラメータ演算ステップで求められた声道フィルタ係数及び前記残差量子化方式決定手段によって決定された量子化方式に基づいて、前記読込みステップで読み込まれた前記素片データを符号化する符号化ステップと、

前記符号化手段ステップで符号化された素片データを前記素片辞書に記録する符号化データ記録ステップとを有することを特徴とする。

【 0 0 1 1 】

上記目的を達成するために本発明の音声合成装置は以下のような構成を備える。即ち、

音声合成用の素片データが記録された音声合成用素片辞書における、所望の素片データが記録された位置を求める素片位置検索手段と、

素片データを復号化するための情報を求め、前記素片辞書から読み出された符号化された素片データを前記情報に従って復号する復号手段と、

前記復号手段で復号された素片データを接続して音声を作成する音声出力手段とを有することを特徴とする。

【0012】

上記目的を達成するために本発明の音声合成方法は以下のような工程を備える。即ち、

音声合成用の素片データが記録された音声合成用素片辞書における、所望の素片データが記録された位置を得る素片位置検索ステップと、

前記素片データに対応する線形予測係数を読み込む予測係数読み込みステップと、

前記素片データに対応する予測残差復号方式を得る予測残差復号方式取得ステップと、

予測符号化された素片データを読み込む予測符号化データ読み込みステップと、

前記読み込みステップで読み込まれた予測符号化データを前記線形予測係数および前記予測残差復号方式に従って復号する予測復号ステップと、

前記予測復号ステップで予測復号された素片データを接続して音声を作成する音声出力ステップとを有することを特徴とする。

また本発明の音声合成方法は以下のような工程を備える。即ち、

音声合成用の素片データが記録された音声合成用素片辞書における、所望の素片データが記録された位置を得る素片位置検索手段と、

前記素片データに対応する声道フィルタ係数を読み込む係数読み込み手段と、

前記素片データに対応する残差復号方式を得る残差復号方式取得手段と、

符号化された素片データを読み込む符号化データ読み込み手段と、

前記読み込み手段によって読み込まれた符号化データを前記声道フィルタ係数および前記残差復号方式に従って復号する復号手段と、

前記復号された素片データを接続して音声を作成する音声出力手段とを有することを特徴とする。

また本発明の音声合成方法は以下のような工程を備える。即ち、

音声合成用の素片データが記録された音声合成用素片辞書における、所望の素

片データが記録された位置を得る素片位置検索ステップと、 前記素片データに対応する声道フィルタ係数を読込む係数読込みステップと、 前記素片データに対応する残差復号方式を得る残差復号方式取得ステップと、 符号化された素片データを読込む符号化データ読込みステップと、

前記読込みステップで読み込まれた符号化データを前記声道フィルタ係数および前記残差復号方式に従って復号する復号ステップと、

前記復号された素片データを接続して音声合成する音声出力ステップと、
を有することを特徴とする。

【 0 0 1 3 】

【発明の実施の形態】

以下、添付図面を参照して本発明の好適な実施の形態を詳細に説明する。本実施の形態に係る音声合成処理は、（１）素片辞書作成、（２）素片辞書を用いた音声合成の２段階に分けられる。従って、以下の説明では、上記２段階の各々について説明する。

【 0 0 1 4 】

【実施の形態 1】

（素片辞書作成）

図 1 は、本発明の実施の形態に係る音声合成装置の概略機能構成を示すブロック図である。

【 0 0 1 5 】

図 1 において、1 0 0 は中央処理装置（CPU）で、数値演算及び各種制御処理を実行し、バス 1 0 5 を介して接続されている後述する各部の動作を制御している。1 0 1 は RAM 及び ROM 等を備えた記憶装置で、中央処理装置 1 0 0 により実行される各種制御プログラムやデータ等を記憶するとともに、中央処理装置 1 0 0 による制御処理で必要となる各種データを一時的に格納している。1 0 2 はハードディスク装置等の外部記憶装置で、素片辞書 1 1 2 および、この素片辞書 1 1 2 の基となる素片データ 1 1 1 が格納されている。1 0 3 は出力装置で、例えばプログラムの動作状況などを表示・出力するユーザインターフェースを有し、更には、音声合成された音声出力するスピーカ等をも備えている。1 0

4 はキーボードやマウス等を備える等の入力装置であり、素片辞書 112 の作成の開始といった制御を行うためのユーザインターフェース機能をも実現している。

【0016】

以上の構成を基に、本実施の形態 1 に係る音声合成装置における動作を説明する。

【0017】

図 3 は、本発明の実施の形態 1 に係る音声合成装置における素片辞書作成処理を示すフローチャートで、この処理を実行するプログラムは記憶装置 101 に記憶されており、中央処理装置 100 の制御の下に実行される。

【0018】

まずステップ S11 で、外部記憶装置 102 に記憶されている素片データ 111 のそれぞれを指示するインデックス i を“0”に初期化する。尚、このインデックス i は記憶装置 101 に設けられている。次にステップ S12 に進み、そのインデックス i で指示される i 番目の素片データ（波形） W_i を読み込む。ここでは、その読み込まれたデータ W_i を、

$$W_i = \{x_0, x_1, \dots, x_{T-1}\}$$

とする。ここで、 T は W_i の時間長（単位はサンプル）である。

【0019】

次にステップ S13 に進み、ステップ S12 で読み込まれた素片データ W_i を 7 ビット μ -law 符号化を用いて符号化する。こうして符号化された結果を、

$$C_i = \{c_0, c_1, \dots, c_{T-1}\}$$

とする。

【0020】

次にステップ S14 に進み、ステップ S13 の 7 ビット μ -law 符号化処理によって生じた符号化歪み ρ を計算する。ここでは、この符号化歪み尺度として、例えば、平均 2 乗誤差

[数 1]

$$\rho = (1/T) \cdot \sum (x_t - \mu(7)^{-1}(c_t))^2$$

を用いることが出来る。ここで、 $\mu(7)^{-1}()$ は、7ビット μ -law復号化関数を示している。なお、ここで、「 2 」は2乗を示し、「 Σ 」は $t=0$ から $t=T-1$ までの総和を示している。

【0021】

次にステップS15に進み、ステップS14の符号化歪み計算で求められた符号化歪み ρ が、予め定められた閾値 ρ_0 より大きいかどうか判定する。もし、 $\rho > \rho_0$ であれば、ステップS16の8ビット μ -law符号化処理に進み、そうでなければステップS17の符号化情報書き込み処理に進む。尚、ステップS16の8ビット μ -law符号化処理では、ステップS12で読み込まれた素片データ W_i を8ビット μ -law符号化を用いて符号化する。こうして符号化された結果を、

$$C_i = \{c_0, c_1, \dots, c_{T-1}\}$$

とする。こうしてステップS17では、その符号化に関する情報（符号化情報）を素片辞書112に書き込む。本実施の形態では、この符号化に関する情報は、符号列 C_i が7ビット μ -law符号化処理の結果であるのか、8ビット μ -law符号化処理の結果であるのかという情報である。従って、符号化情報として、

0 ... 7ビット μ -law符号化時

1 ... 8ビット μ -law符号化時

を書き込む。

【0022】

次にステップS18に進み、こうして符号化された素片データ C_i を素片辞書112に書き込む。そしてステップS19に進み、インデックス i の値を基に、 N 個の素片データ全てについて処理が終わっているかを判定する。ここで $i=N-1$ が成立すれば、本実施の形態の処理を終了し、そうでなければステップS20に進んでインデックス i を+1して更新し、再びステップS12に進んで、そのインデックス i で指示される素片データの読み込みを行い、前述の処理を繰返す。

【0023】

(音声合成)

図4は、本実施の形態1に係る音声合成装置における素片辞書の検索処理を示

すフローチャートで、この処理は前述した図2のステップS4の処理を詳細化したサブフローチャートとなっている。尚、この処理を実行する制御プログラムも記憶装置101に記憶されている。

【0024】

図4において、まずステップS21で、読込むべき素片データの位置を検索する。具体的には、素片辞書112が外部記憶装置（ハードディスク）102に配置されているならば、読込むべき素片データの位置をシークする。一方、もし素片辞書112がRAM等のメモリ上に配置されているならば、読込むべき素片データの位置へポインタ（アドレスレジスタ）を移動することになる。次にステップS22に進み、前述した符号化に関する情報（符号化情報）を読込む。本実施の形態では、この符号化情報は、7ビット μ -lawを用いているのか、或は8ビット μ -law用いているのかという情報であり、

符号化情報が“0”であれば、7ビット μ -law

符号化情報が“1”であれば、8ビット μ -law

である。

【0025】

次にステップS23で、 μ -law符号化のビット数を判定し、もし符号化情報が7ビット μ -law符号化（=0）を示しているならばステップS24に進み、7ビット μ -law符号化データの読込み処理を行い、符号化情報が8ビット μ -law符号化（=1）を示しているならばステップS26に進んで、8ビット μ -law符号化データの読込み処理を行う。ステップS24では、7ビット μ -law符号化された素片データを読込んでステップS25に進み、その読み込んだ7ビット μ -law符号化された素片データを復号化し、この素片辞書の検索処理を終了する。

【0026】

一方、ステップS26では、8ビット μ -law符号化された素片データを読込んでステップS27に進み、その8ビット μ -law符号化された素片データを復号化して、この素片辞書の検索処理を終了する。

【0027】

以上説明したように本実施の形態1によれば、素片データを7ビット μ -law符

号化により符号化し、そのときの符号化歪みが所定値（ ρ_0 ）以上であれば 8 ビット μ -law 符号化に切り換えて符号化することにより、演算量を少なくして、充分な圧縮性能を得ることができるという効果がある。

【0028】

〔実施の形態 2〕

次に、素片データ毎に量子化テーブルを変更する本発明の実施の形態 2 について説明する。

【0029】

（素片辞書作成）

図 5 は、本発明の実施の形態 2 に係る素片辞書 112 の作成処理を示すフローチャートで、この処理を実行するプログラムは記憶装置 101 に記憶されている。以下、このフローチャートに即して説明を行う。

【0030】

まずステップ S31 で、外部記憶装置 102 に記憶された素片データ 111 を指示するインデックス i を“0”に初期化する。尚、このインデックス i は記憶装置 101 に設定されている。次にステップ S32 に進み、そのインデックス i で指示される i 番目の素片データ（波形） W_i を読み込む。こうして読み込まれたデータを、

$$W_i = \{x_0, x_1, \dots, x_{T-1}\}$$

とする。ここで、 T は W_i の時間長（単位はサンプル）である。

【0031】

次にステップ S33 に進み、ステップ S32 で読み込んだ素片データ W_i に対するスカラ量子化コードブック Q_i を作成する。このコードブック Q_i は、素片データ W_i 及びスカラ量子化コードブック Q_i を用いて素片データ W_i を符号化した後、復号化したデータ系列

$$Y_i = \{y_0, y_1, \dots, y_{T-1}\}$$

について、平均 2 乗誤差

〔数 2〕

$$(1/T) \cdot \sum (x_t - y_t)^2$$

が最小となるように設計する。具体的には、LBG法などのアルゴリズムを用いることが出来る。なお、ここで、「 \cdot^2 」は2乗を示し、「 Σ 」は $t=0$ から $t=T-1$ までの総和を示している。

【0032】

次にステップS34に進み、ステップS33で作成したスカラ量子化コードブック Q_i を素片辞書112に書き込む。そしてステップS35で、その作成した量子化コードブック Q_i を用いて、素片データ W_i を符号化（スカラ量子化）する。

【0033】

ここでコードブック Q_i を

$$Q_i = \{q_0, q_1, \dots, q_{N-1}\} \quad (N \text{ は量子化ステップ})$$

とすると、 $x_t (\in W_i)$ に対応する符号 c_t は、

[数3]

$$c_t = \arg \min_n (x_t - q_n)^2 \quad (0 \leq n < N)$$

となる。

【0034】

次にステップS36に進み、ステップS35で符号化された素片データ $C_i (= \{c_0, c_1, \dots, c_{T-1}\})$ を素片辞書112に書き込む。次にステップS37に進み、N個の素片データ全てについて処理が終わっているかを判定する。もし $i = N-1$ が成立すれば本実施の形態の処理を終了し、そうでなければステップS38でインデックス i を+1してステップS32に戻り、その更新されたインデックス i に基づいて、次の素片データを読込む。このような処理をN個の素片データを全てに対する処理が終了するまで繰り返し実行する。

【0035】

(音声合成)

次に、図5の素片辞書作成手順によって作成した素片辞書112を用いた音声合成方法について、図6のフローチャートを参照して説明する。

【0036】

図6は本発明の実施の形態2に係る音声合成処理を示すフローチャートである

【0037】

まずステップS41で、読み込むべき素片データの位置を検索する。具体的には、素片辞書112が外部記憶装置102などのハードディスクに配置されているならば読み込むべき素片データの位置をシークし、また素片辞書112がRAM等のメモリ上に配置されているならば、読み込むべき素片データの位置へポインタ（アドレスレジスタ）を移動することに相当する。次にステップS42に進み、スカラ量子化コードブックQiを読み込む。そしてステップS43に進み、符号化された素片データCiを読み込む。そしてステップS44で、ステップS42で読み込んだスカラ量子化コードブックQiを用いて、ステップS43で読み込んだ符号化された素片データCiを復号化する。これにより、素片辞書検索のサブフローを終了する。

【0038】

これにより本実施の形態2によれば、素片データをスカラ量子化コードをブックを用いてスカラ量子化（符号化）し、その符号化に使用した量子化コードブックとともに素片辞書に記憶することにより、演算量を少なくし、圧縮効率を高めて記憶することができる。

【0039】

〔実施の形態2の変形〕

この実施の形態2において、前述の実施の形態1と同様に、波形1サンプル当りのビット数、即ち、スカラ量子化の量子化ステップ数を素片データ毎に変更することも可能である。これは、図5におけるステップS34のコードブック書き込み処理に先立って、量子化ステップ数を素片辞書112に記録する。そして音声合成の際には、ステップS42のスカラ量子化コードブックの読み込み処理に先立って、量子化ステップ数を素片辞書112から読み込むことによって実現できる。この量子化ステップ数の決定には、前述の実施の形態1と同様に量子化歪みを基準として用いることができる。

【0040】

〔実施の形態3〕

前述の実施の形態 2 では、素片データ 111 とスカラ量子化コードブックは 1 対 1 に対応しているが、複数の素片データに対して 1 つのスカラ量子化コードブックを設計しても良い。この場合の一実施の形態を以下に示す。

【0041】

(素片辞書作成)

図 7 は、本発明の実施の形態 3 に係る素片辞書作成処理を示すフローチャートである。以下、このフローチャートに即して説明を行う。

【0042】

まずステップ S51 で、素片辞書 112 に含まれる全ての素片データを読み込む。次にステップ S52 に進み、これら全ての素片データを複数 (M 個とする) のクラスタにクラスタリングする。このクラスタリングの手法として、例えば LBG アルゴリズムを用いることができる、この際、任意の 2 素片データ間に距離を定義する必要がある。そこで、例えば、各素片データをケプストラムなどのパラメータに変換し、2 つの素片データ間の DP マッチングを取ることで距離を定めることが可能である。

【0043】

次にステップ S53 で、素片クラスタを指示するインデックス i を “0” に初期化する。次にステップ S54 に進み、素片クラスタ L_i に対するスカラ量子化コードブック Q_i を作成する。次にステップ S55 に進み、ステップ S54 のスカラ量子化コードブック作成処理で作成したコードブック Q_i を素片辞書 112 に書き込む。次にステップ S56 に進み、M 個の素片クラスタの全てについて処理が終わっているかを判定する。もし $i = M - 1$ が成立すれば (M 個の素片クラスタの全てについて処理が終了) ステップ S58 に進み、インデックス i を “0” に初期化し、そうでなければステップ S57 に進んで、クラスタインデックス i に 1 を加え、再びステップ S54 のスカラ量子化コードブック作成処理に戻る。

【0044】

M 個の素片クラスタの全てについて処理が終了した後ステップ S58 で、素片データを指示するインデックス i を “0” に初期化する。次にステップ S59 に

進み、素片データ W_i に対するスカラ量子化コードブック Q_i を選択する。ここではスカラ量子化コードブック Q_i として、ステップ S 5 2 の素片データクラスタリングの際に、素片データ W_i が属していたクラスタに対応する量子化コードブックを選ぶことができる。

【 0 0 4 5 】

次にステップ S 6 0 に進み、ステップ S 5 9 で選択したスカラ量子化コードブックを指示する情報を素片辞書 1 1 2 に書き込む。そしてステップ S 6 1 で、ステップ S 5 4 で作成したコードブック Q_i を用いて素片データ W_i を符号化する。そしてステップ S 6 2 に進み、その符号化された素片データ $C_i (= \{c_0, c_1, \dots, c_{T-1}\})$ を素片辞書 1 1 2 に書き込む。次にステップ S 6 3 に進み、N 個の素片データ全てについて処理が終わっているかを判定する。もし $i = N - 1$ が成立すれば本実施の形態の処理を終了し、そうでなければステップ S 6 4 に進み、素片データのインデックス i に “1” を加えてステップ S 5 9 に戻り、前述の処理を実行する。

【 0 0 4 6 】

(音声合成)

次に本実施の形態 3 に係る、前述の素片辞書作成手順によって作成した素片辞書を用いた音声合成方法を、図 9 のフローチャートを参照して説明する。ここでは説明を簡単にするため、全素片クラスタに対応するコードブックが全ての処理(図 2)に先立って記憶装置 1 0 1 に予め読込まれているものとする。

【 0 0 4 7 】

まずステップ S 7 1 で、読込むべき素片データの位置を検索する。具体的には、素片辞書 1 1 2 が外部記憶装置 1 0 2 などのハードディスクに配置されているならば、読込むべき素片データの位置をシークし、また素片辞書 1 1 2 が RAM 等のメモリ上に配置されているならば、読込むべき素片データの位置へポインタ(アドレスレジスタ)を移動する。次にステップ S 7 2 に進み、使用すべきスカラ量子化コードブック Q を決定するための情報を読込む。そしてステップ S 7 3 に進み、スカラ量子化コードブック Q を得る。

【 0 0 4 8 】

次にステップ S 7 4 に進み、符号化された素片データ C を読込む。そしてステップ S 7 5 で、ステップ S 7 3 のコードブック検索で得られたスカラ量子化コードブック Q を用いて、符号化された素片データ C を復号化し、素片辞書検索のサブフローを終了する。

【0049】

これにより本実施の形態 3 によれば、複数の素片データに対してスカラ量子化コードブックを適用してスカラ量子化（符号化）し、その符号化に使用した量子化コードブック情報とともに素片辞書に記憶することにより、演算量を少なくし、圧縮効率を高めて記憶することができる。

【0050】

〔実施の形態 3 の変形 1〕

前述の実施の形態 3 において、ステップ S 5 2（図 7）の素片データのクラスタリングにおいて、通常のクラスタリング手法を用いる代りに、素片データ 1 1 1 に付与されたラベル（音素，VCV，etc.）を用いることも可能である。この場合、ステップ S 7 3（図 9）のコードブックの検索処理で必要となる情報は、ステップ S 7 1 の素片データ位置検索のための情報に含まれることとなり、ステップ S 7 2 のコードブック情報の読み込み処理は不要になる。

【0051】

〔実施の形態 3 の変形 2〕

前述の実施の形態 3 において、ステップ S 5 9 のスカラ量子化コードブックの選択処理（図 8）において、ステップ S 5 4 のスカラ量子化コードブック作成ステップで作成された全てのコードブックの中から、性能の最も良い（量子化歪みが最小となる）コードブックを選んでも良い。

【0052】

〔実施の形態 4〕

次に、素片データ毎に線形予測係数を設定して予測誤差を符号化する本発明の実施の形態 4 について述べる。

【0053】

（素片辞書作成）

図10は、本発明の実施の形態4に係る素片辞書112の作成処理を示すフローチャートである。以下、本フローチャートに即して説明を行う。

【0054】

まずステップS81で、素片データ111を指示するインデックス*i*を“0”に初期化する。次にステップS82で、インデックス*i*で指示される*i*番目の素片データ（波形）*W_i*を読み込む。この読み込まれたデータを、

$$W_i = \{x_0, x_1, \dots, x_{T-1}\}$$

とする。ここで、*T*は*W_i*の時間長（単位はサンプル）である。

【0055】

次にステップS83に進み、素片データ*W_i*に対する線形予測係数および予測残差を計算する。ここで線形予測次数を*L*次とすると、この線形予測モデルは予測係数*a_l*および予測残差*d_t*を用いて、以下のように表される。

[数4]

$$x_t = \sum a_l x_{t-l} + d_t$$

ここで \sum は*l* = 1 ~ *L*までの総和を示している。

【0056】

そこで、予測残差二乗和、

[数5]

$$\sum d_t^2$$

を最小化するように予測係数を決定する。ここで \sum は*t* = 1 ~ *T* - 1までの総和を示し、*d_t*²は「*d_t*」の2乗を示している。

【0057】

次にステップS84に進み、ステップS83の線形予測係数、予測残差計算で計算された予測係数を素片辞書112に書き込む。そしてステップS85に進み、ステップS83の線形予測係数、予測残差計算で計算された予測残差*d_t*に対する予測残差量子化コードブック*Q_i*を作成する。この予測残差量子化コードブック*Q_i*は、予測残差*d_t*およびコードブック*Q_i*を用いて*d_t*を符号化した後、復号化したデータ系列

$$E_i = \{e_1, e_{l+1}, \dots, e_{T-1}\}$$

について、平均 2 乗誤差

[数 6]

$$(1/T) \cdot \sum (d_t - e_t)^2$$

が最小となるように設計する。ここで \sum は $t = 1 \sim T-1$ までの総和を示している。この設計は、具体的には、LBG 法などのアルゴリズムを用いることが出来る。

【0058】

次にステップ S 8 6 に進み、ステップ S 8 5 の予測残差量子化コードブック作成処理で作成されたコードブック Q_i を素片辞書 1 1 2 に書き込む。そしてステップ S 8 7 に進み、ステップ S 8 3 で計算された線形予測係数 a_l 、及びステップ S 8 5 で作成したコードブック Q_i を用いて素片データ W_i を符号化する。ここではコードブック Q_i を、

$$Q_i = \{q_0, q_1, \dots, q_{N-1}\} \quad (N \text{ は量子化ステップ})$$

とすると、 $x_t (\in W_i)$ に対応する符号 c_t は、

[数 7]

$$c_t = \arg \min_n (x_t - \sum a_l y_{t-l} - q_n)^2 \quad (0 \leq n < N)$$

となる。ここで、 y_t は x_t を本手法で符号化した後、復号した値である。

【0059】

次にステップ S 8 8 に進み、その符号化された素片データ $C_i (= \{c_0, c_1, \dots, c_{T-1}\})$ を素片辞書 1 1 2 に書き込む。そしてステップ S 8 9 で、 N 個の素片データ全てについて処理が終わっているかを判定する。もし $i = N-1$ が成立すれば本実施の形態 4 に係る素片辞書の作成処理を終了し、そうでなければステップ S 9 0 に進んで、素片データのインデックス i に 1 を加え、再びステップ S 8 2 の素片データ読み込み処理に戻る。

【0060】

(音声合成)

次に、図 10 の素片辞書の作成手順によって作成した素片辞書 1 1 2 を用いた音声合成方法について図 11 のフローチャートを参照して説明する。

【0061】

まずステップS101で、読込むべき素片データの位置を検索する。具体的には、素片辞書112がハードディスク上に配置されているならば、読込むべき素片の位置をシークし、また素片辞書112がRAM等のメモリに配置されているならば、読込むべき素片の位置へポインタ（アドレスレジスタ）を移動することになる。

【0062】

次にステップS102に進み、予測係数 a_i を読込む。そしてステップS103に進み、予測残差量子化コードブック Q_i を読込む。次にステップS104に進み、予測残差データ C_i を読込む。そしてステップS105で、ステップS102で読込まれた予測係数、及びステップS103で読込まれた予測残差量子化コードブック、及び復号化すべきサンプルの直前のLサンプルの復号後データを用いて、符号化された素片データ C_i を復号化し、素片辞書検索のサブフローを終了する。

【0063】

以上説明したように本実施の形態4によれば、素片データを線形予測し、その予測残差を予測残差量子化コードブックを用いて予測符号化して素片辞書を作成し、その素片データの再生の際には、その予測係数、コードブック、予測残差を読み込んで素片データを復号することにより、演算量を少なくして、十分な圧縮性能を得ることができるという効果がある。

【0064】

〔実施の形態4の変形1〕

実施の形態4において、前述の実施の形態1と同様に、波形データ1サンプル当りのビット数、即ち、スカラ量子化の量子化ステップ数を素片データ毎に変更することも可能である。

【0065】

これは、図5のステップS34のコードブック書き込み処理に先立って、量子化ステップ数を素片辞書112に記録し、音声合成の際には、ステップS103の予測残差コードブック読み込み処理に先立って量子化ステップ数を素片辞書112から読込むことによって実現できる。この量子化ステップ数の決定には、前述

の実施の形態 1 と同様に、例えば量子化歪み ρ を基準として用いることができる。

【0066】

〔実施の形態 4 の変形 2〕

また本実施の形態 4 において、素片データ毎に線形予測次数を変更することも可能である。これはステップ S 8 4 の予測係数書き込み処理に先立って、予測次数を素片辞書 1 1 2 に記録し、音声合成の際には、ステップ S 8 8 の予測係数の読み込み処理に先立って予測次数を素片辞書 1 1 2 から読み込むことにより実現できる。この予測次数の決定には、量子化歪みを基準として用いることができる。

【0067】

〔実施の形態 4 の変形 3〕

本実施の形態 4 において、ステップ S 8 5 の予測残差量子化コードブック作成処理の後、予測残差量子化コードブック更新を行うことにより、符号化性能を向上させることが可能である。これはステップ S 8 5 の予測残差量子化コードブック作成処理では、予測残差 d_t に対してコードブックを最適化していたのに対し、ステップ S 8 7 の予測符号化処理では、

〔数 8〕

$$x_t - \sum a_l y_{t-l} \quad (\neq d_t = x_t - \sum a_l x_{t-l})$$

について、コードブック参照が行われるためである。この量子化コードブックを更新するアルゴリズムとしては、A b S (Analysis by Synthesis) 法等を用いることができる。ここで Σ は、 $l = 1 \sim L$ の総和を示している。

【0068】

〔実施の形態 4 の変形 4〕

前述の実施の形態 4 では、素片データと予測残差量子化コードブックは 1 対 1 に対応しているが、複数の素片データに対して予測残差量子化コードブックを設計しても良い。この具体的な方法は、前述の実施の形態 2 に対する実施の形態 3 への変形と同じである。

【0069】

〔実施の形態 4 の変形 5〕

前述の実施の形態 4 において、素片データの開始から L サンプルのデータについては符号化を行わず、そのまま素片辞書 1 1 2 に書き込んでも良い。これにより、素片データの開始から L サンプルについては線形予測がうまく行われないう現象を回避することが可能である。

【0 0 7 0】

〔その他の実施の形態〕

前述の実施の形態 2, 3, 4 では、量子化の手法としてスカラ量子化を用いたが、連続する複数のサンプルを 1 つのベクトルと見做し、ベクトル量子化を適用することも可能である。

【0 0 7 1】

また、前述した各実施の形態は、それぞれ独立した方法として例示されているが、任意の組合わせで併用することが可能である。例えば、素片データ毎に実施の形態 2 の符号化と実施の形態 4 の符号化を切り替えて用いるようにしてもよい。この場合、符号化方法の選択基準として、前述の実施の形態 1 で用いた符号化歪みによる方法を用いることができる。同様に従来例に示したような既存の符号化方法とも併用することが可能である。

【0 0 7 2】

また、破裂音の様な非定常性の素片については、破裂前・破裂後といった様に素片データを分割することによって、符号化性能をより向上させることも可能である。

また上記実施の形態 4 では、線形予測モデルに基づいて説明を行ったが、その他の声道フィルタモデルに本発明を適用することも可能である。その場合のモデルパラメータ演算は、例えば線形予測係数の代わりに LMA (Log Magnitude Approximation) フィルタ係数を用い、予測残差の代わりに LMA フィルタの残差を用いることによって、本発明をケプストラム方に適用することができる。この実施の形態は本願請求項 2 1 乃至 2 4 に相当するものである。

【0 0 7 3】

なお本発明は、複数の機器（例えばホストコンピュータ、インターフェース機器、リーダ、プリンタなど）から構成されるシステムに適用しても、一つの機器

からなる装置（例えば、複写機、ファクシミリ装置など）に適用してもよい。

【0074】

また、本発明の目的は、前述した実施形態の機能を実現するソフトウェアのプログラムコードを記録した記憶媒体（または記録媒体）を、システムあるいは装置に供給し、そのシステムあるいは装置のコンピュータ（またはCPUやMPU）が記憶媒体に格納されたプログラムコードを読み出し実行することによっても達成される。この場合、記憶媒体から読み出されたプログラムコード自体が前述した実施形態の機能を実現することになり、そのプログラムコードを記憶した記憶媒体は本発明を構成することになる。また、コンピュータが読み出したプログラムコードを実行することにより、前述した実施形態の機能が実現されるだけでなく、そのプログラムコードの指示に基づき、コンピュータ上で稼働しているオペレーティングシステム(OS)などが実際の処理の一部または全部を行い、その処理によって前述した実施形態の機能が実現される場合も含まれる。

【0075】

さらに、記憶媒体から読み出されたプログラムコードが、コンピュータに挿入された機能拡張カードやコンピュータに接続された機能拡張ユニットに備わるメモリに書込まれた後、そのプログラムコードの指示に基づき、その機能拡張カードや機能拡張ユニットに備わるCPUなどが実際の処理の一部または全部を行い、その処理によって前述した実施形態の機能が実現される場合も含まれる。

【0076】

以上説明したように本実施の形態によれば、素片データに応じて符号化方式を選択することによって高い符号化性能を得ることができ、また復号の際の適応アルゴリズムを廃することによって波形処理に基く音声合成の演算量の少なさを生かすことが可能となる。

【0077】

【発明の効果】

以上説明したように本発明によれば、演算量を少なく抑え、かつ十分な圧縮性能が得られる素片データの符号化を行って素片辞書を作成できるという効果がある。

【 0 0 7 8 】

また本発明によれば、作成された素片辞書を用いて音声合成できる音声合成方法及び装置を提供することができるという効果がある。

【図面の簡単な説明】

【図 1】

本発明の実施の形態 1 に係る音声合成装置のハードウェア構成を示すブロック図である。

【図 2】

一般的な音声合成処理を示すフローチャートである。

【図 3】

本発明の実施の形態 1 に係る音声合成装置における素片辞書の作成処理を示すフローチャートである。

【図 4】

本発明の実施の形態 1 に係る音声合成装置における素片辞書の検索処理を示すフローチャートである。

【図 5】

本発明の実施の形態 2 に係る素片辞書の作成処理を示すフローチャートである。

【図 6】

本発明の実施の形態 2 に係る素片辞書の検索を行う音声合成処理を示すフローチャートである。

【図 7】

本発明の実施の形態 3 に係る素片辞書作成処理を示すフローチャートである。

【図 8】

本発明の実施の形態 3 に係る素片辞書作成処理を示すフローチャートである。

【図 9】

本発明の実施の形態 3 に係る素片辞書の検索を行う音声合成処理を示すフローチャートである。

【図 1 0】

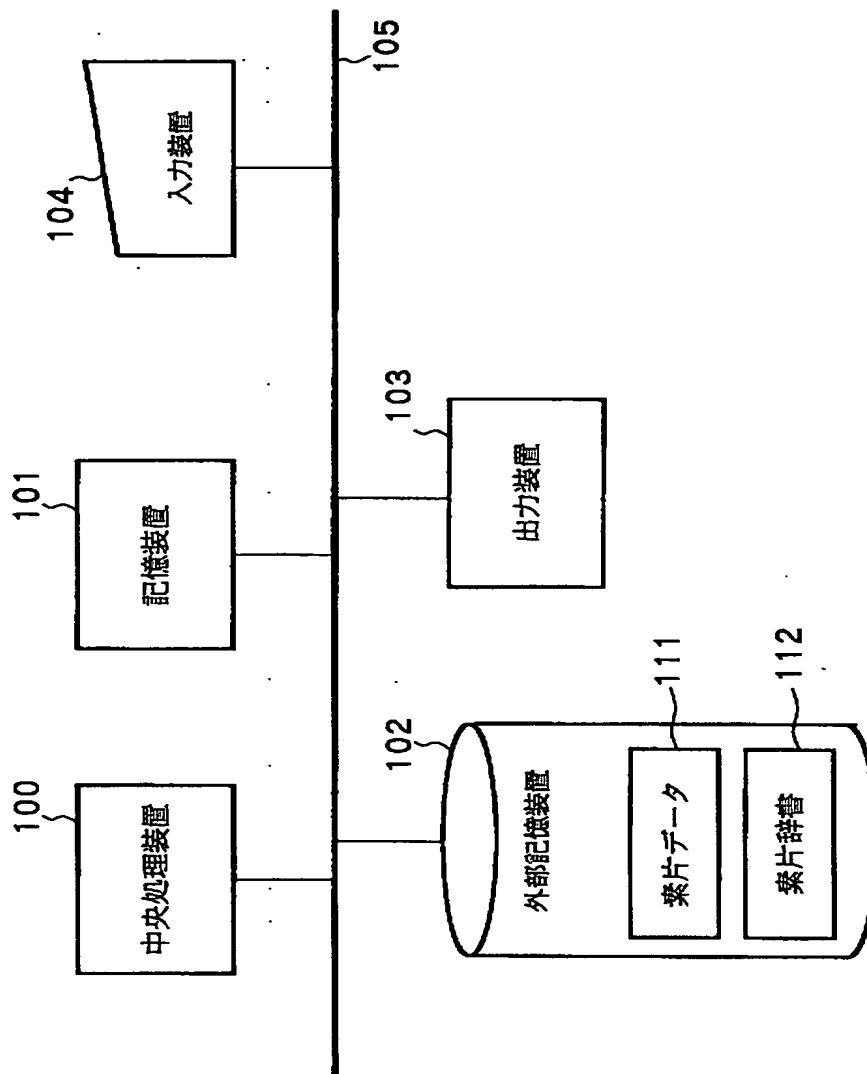
本発明の実施の形態 4 に係る素片辞書の作成処理を示すフローチャートである

【図 1 1】

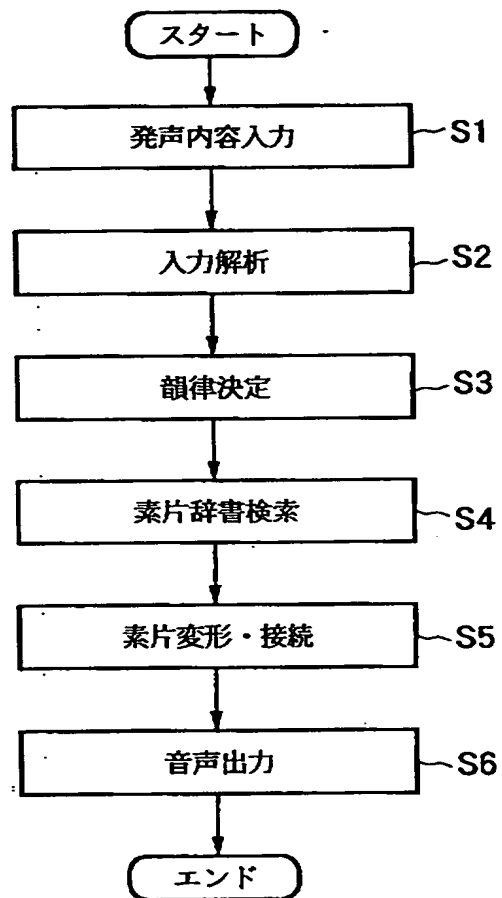
本発明の実施の形態 4 に係る素片辞書の検索、復号を行う音声合成処理を示すフローチャートである。

【書類名】 図面

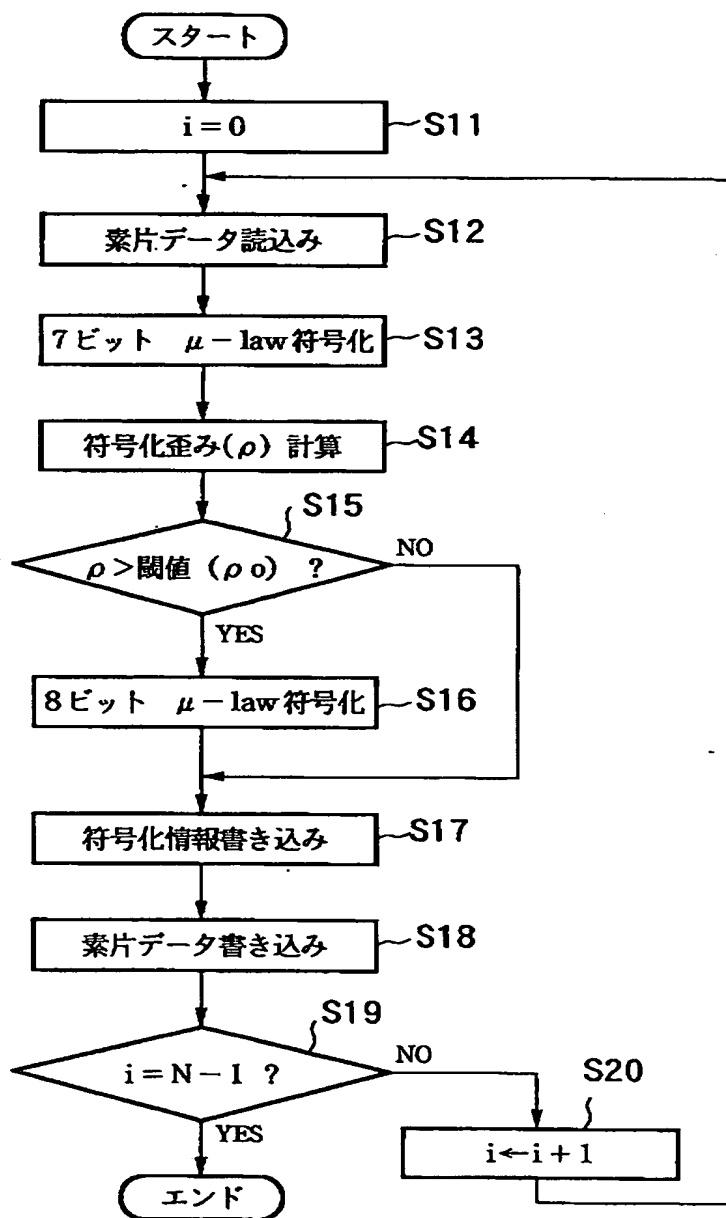
【図 1】



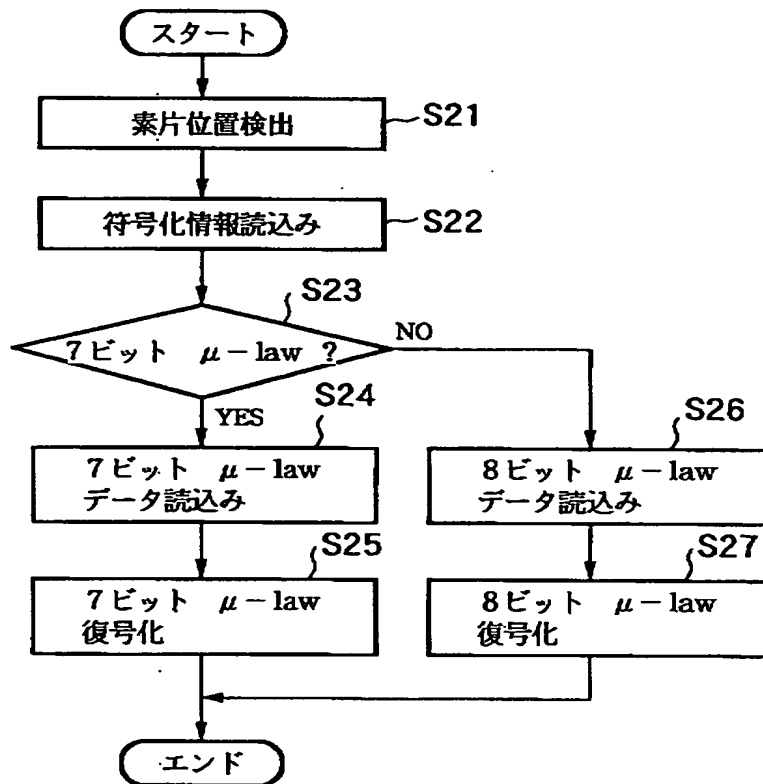
【図 2】



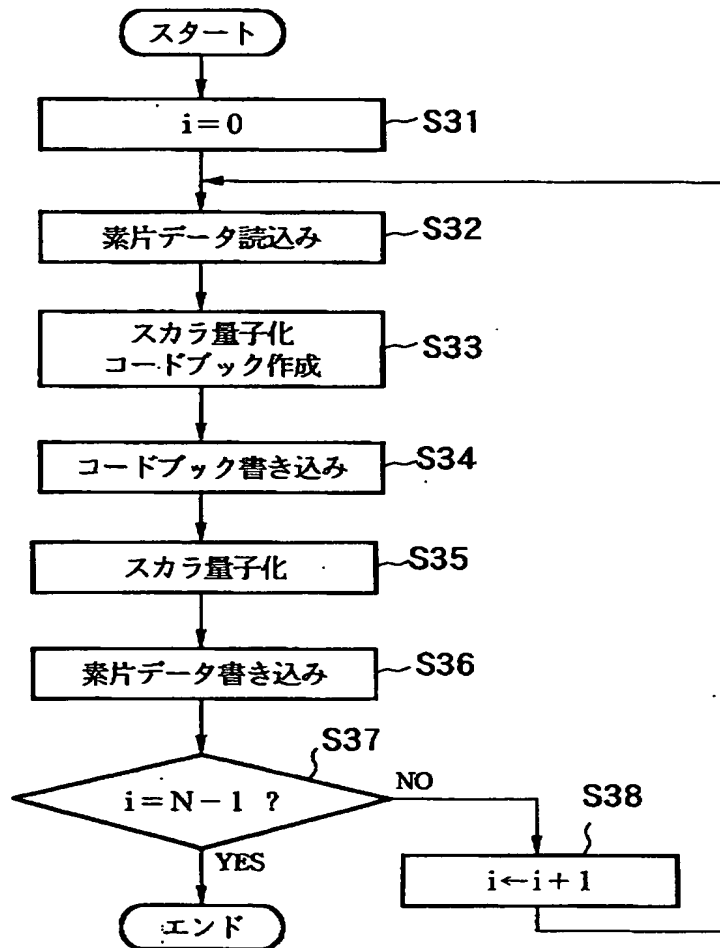
【図 3】



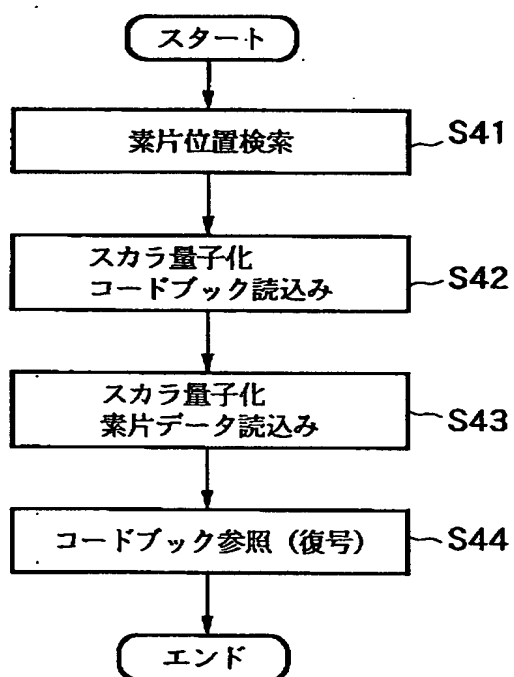
【図 4】



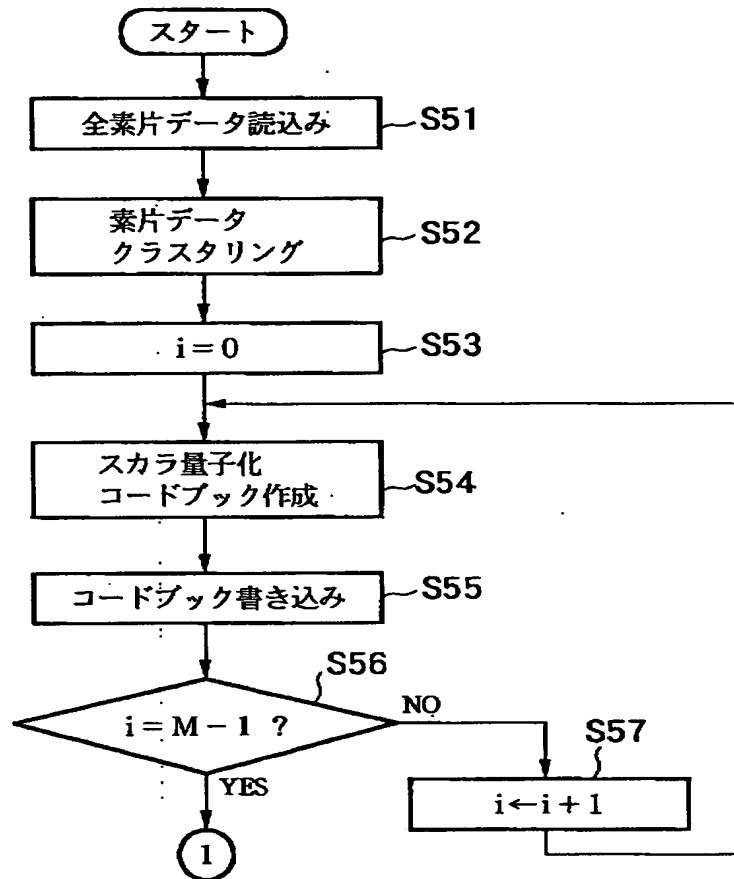
【図 5】



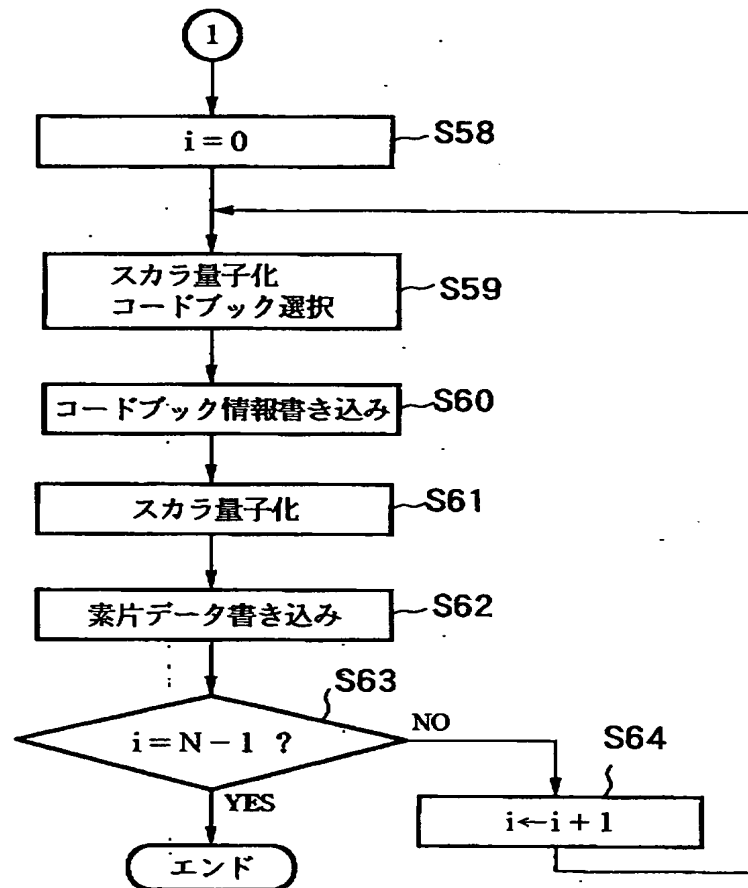
【図 6】



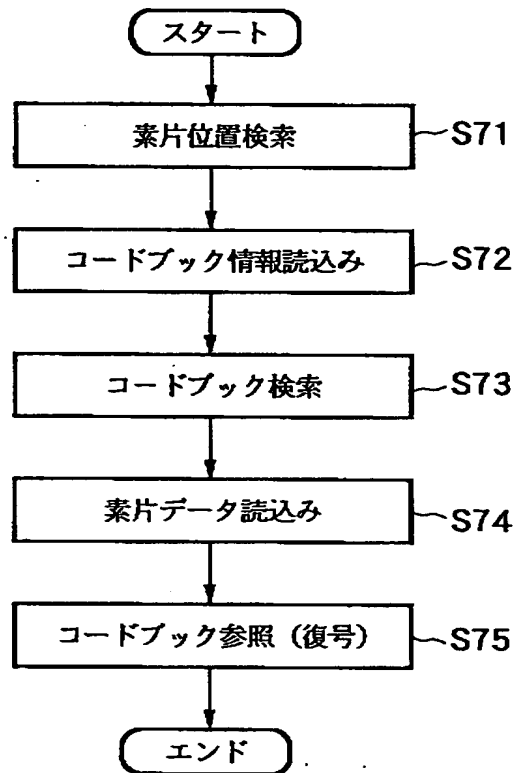
【図 7】



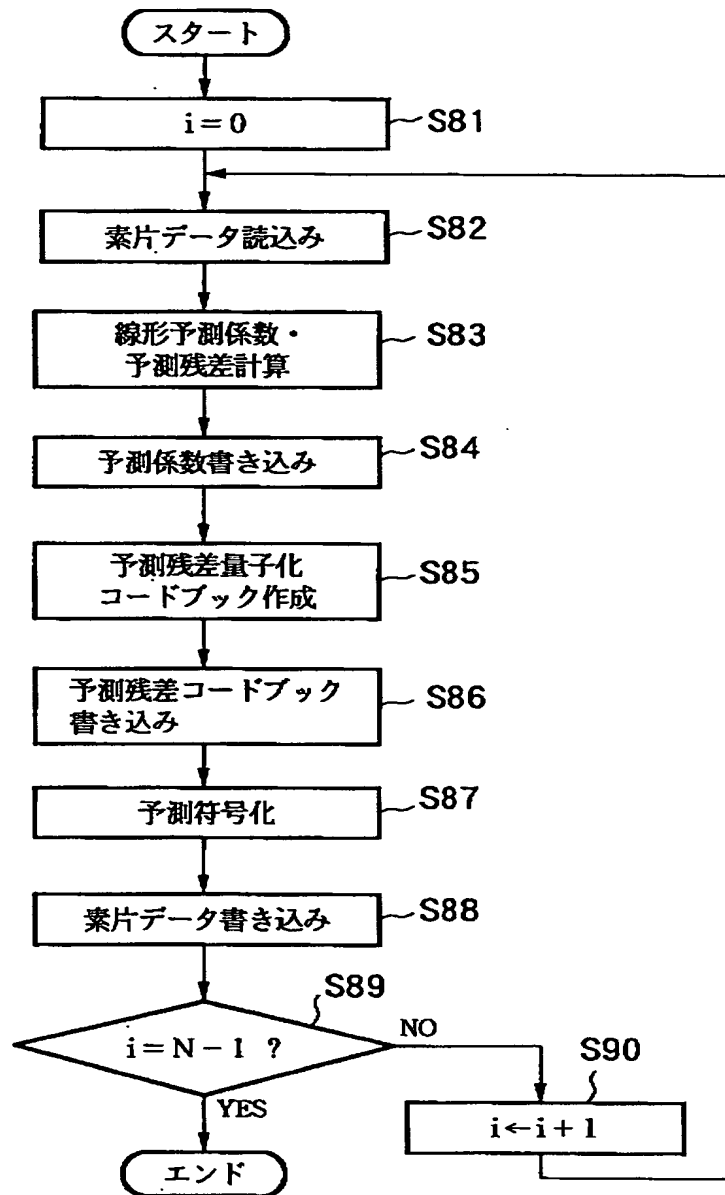
【図 8】



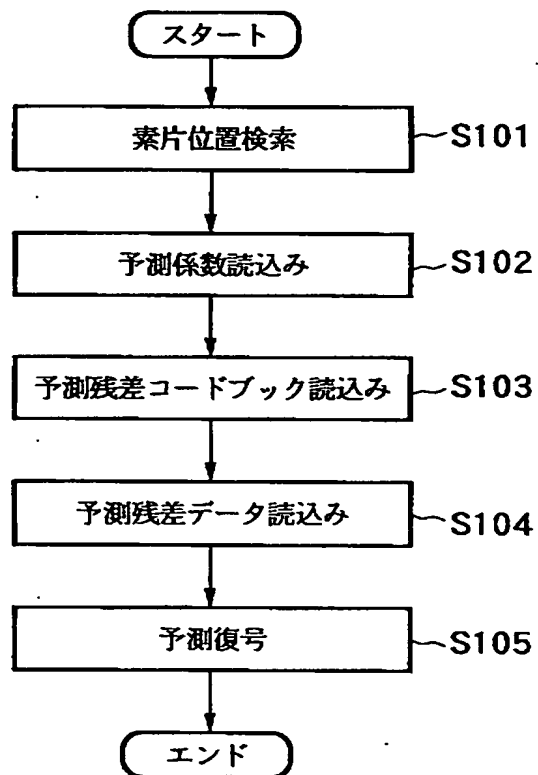
【図 9】



【図 10】



【図 1 1】



【書類名】 要約書

【要約】

【課題】 演算量を少なく抑え、かつ十分な圧縮性能が得られる素片データの符号化を行って素片辞書を作成する。

【解決手段】 音声が記録されたデータベースから合成単位に切り出された素片データを読み込み、その素片データに対して7ビット μ -law符号化を実行し（S 1 3）、その際の符号化歪みを計算する（S 1 4）。その歪みが所定閾値よりも大きい場合には、その素片データを8ビット μ -law符号化する。こうして符号化が、7ビット或は8ビット μ -law符号化のいずれであるかを示す符号化情報とともに、その符号化した素片データを素片辞書に登録する。

【選択図】 図3

出 願 人 履 歴 情 報

識別番号 [000001007]

1. 変更年月日 1990年 8月30日
[変更理由] 新規登録
住 所 東京都大田区下丸子3丁目30番2号
氏 名 キヤノン株式会社